

Columbia Law School

Scholarship Archive

Faculty Scholarship

Faculty Publications

2018

Review: Ralph Wedgwood, The Value of Rationality

Joseph Raz

Columbia Law School, raz@law.columbia.edu

Follow this and additional works at: https://scholarship.law.columbia.edu/faculty_scholarship



Part of the [Law Commons](#)

Recommended Citation

Joseph Raz, *Review: Ralph Wedgwood, The Value of Rationality*, MIND, VOL. 127, P. 1253, 2018 (2018).

Available at: https://scholarship.law.columbia.edu/faculty_scholarship/2462

This Working Paper is brought to you for free and open access by the Faculty Publications at Scholarship Archive. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarship Archive. For more information, please contact scholarshiparchive@law.columbia.edu.

Review: Ralph Wedgwood, *The Value of Rationality*, OUP 2017, 267pp.
By Joseph Raz, Columbia University and KCL

The book is the first of a planned trilogy, with *Rationality and Belief* and *Rationality and Choice* to follow, titles which signal Wedgwood's view that one concept of rationality applies across the practical and epistemic domains. The book's chapters deal with a variety of disparate issues (e.g., 'ought implies can', the nature of reasons, the meaning of 'ought', the structure of virtue, the virtue of rationality, comparative values), each well worth studying for its own sake, and each explored with reference to the history of the subject, and to some of the relevant discussions. Yet, together they combine into a theory of rationality and its value. Wedgwood is developing a theory of the concept in order to understand the property it expresses. The concept is used primarily to evaluate mental states (belief, intention) and events (decision, judgement), an evaluation that is based purely on the degree to which the mental states and events present in the thinker's mind at or shortly before the relevant time (internalism), constitute a coherent pattern. Norms of rationality, i.e. principles specifying what rationality requires, are constitutive of the mental state to which they apply. I assume that that means that what makes the mental state one of that kind is that certain requirements of rationality apply to it, those that apply to the kind. Wedgwood underlines another aspect: each type of mental state and event that can be rational has an aim, and thinking rationally is a means to satisfying the end. Beliefs aim at truth and choice aims at the practicable good. The means to the ends are mental states that are correct, that is, ones that conform to norms of correctness that are themselves constitutive of the types of mental state, though they evaluate mental states by their relations to the external world. Both rationality and correctness come in degrees. The more irrational a person's ways of thinking are the more incorrect his thinking is likely to be.

In my view Wedgwood limits his achievement by misidentifying his topic. "To think rationally is to think as one should" (1), he writes, encouraging the hope that he will conceive of rationality as the virtue of reason, a virtue manifested in the proper exercise of rational powers. However, 'rationality', as Wedgwood understands it, is confined to "the proper use of our reasoning faculties" (32), and important as the reasoning faculties are, they are only part of our rational faculties, which include powers of concentration, attention, decisiveness, discrimination, sensitivity and more.

Equally, if not more damaging to his ambition is his claim that the proper use of our reasoning faculties is judged by the outcome of their use, which leads him to identify rationality with a pattern of mental states and events. In fact, failure to reason properly is manifested in the manner of reasoning and not in its results. One's reasoning may be faulty, even irrational, when the result is fortunate (and, say, displays the right pattern). Reasoning in a reckless, inattentive, negligent etc. manner is irrational whatever the result. However, Wedgwood is right to point out that for the most part studies of rationality in decision theory and formal epistemology are taken to be a study of the coherence of (some) mental states, and one of his aims is to provide an account that applies to those disciplines.

Wedgwood regards as his main challenge to explain the relevance of coherence to rationality. "The norms of rationality are both pervasive ... and have an inescapable grip on us - we necessarily already have a disposition to conform to them. However, it is puzzling how there can be any norms of this sort that evaluate mental events purely on the basis of coherence. Why does coherence matter? Surely it is nothing more than just a pretty pattern of mental states. Why is it a matter of any importance whatsoever whether one's mental

states form this pretty pattern or not?" (4) The reader has to wait to the last couple of chapters for Wedgwood's answer. First, he turns to his thesis that there is only one concept of rationality – a view that can be properly debated only when we confront at least two conflicting accounts of rationality and the question arises whether at least one of them is mistaken or whether perhaps both are correct accounts of different concepts of rationality, both of which have their legitimate uses. Then he turns to the first of several arguments that the concept of rationality is a normative concept: "There is a way of using the terms 'rational' and 'justified' so that phrases like 'rational belief' mean exactly the same thing as 'justified belief' ... in this book, I shall be using the terms 'rational' and 'justified' in this way. But the term 'justified' surely expresses a normative concept." (26-7) More significant is the observation: "To use our faculty of reason 'properly' is to use it as it should be used. When a belief or a choice derives from the proper use of this faculty it thereby has what in the broadest sense could be called a kind of virtue or excellence - a feature in virtue of which beliefs (and the like) can count as good, or worthy of a certain kind of commendation". (25-6) Moreover, "it seems plausible that irrationality is always and necessarily some kind of defect or flaw or blemish ... This makes it plausible that it is a conceptual truth that irrationality is some kind of defect or flaw." (26) So, the concept of rationality is itself a normative concept.

At a minimum these points show that 'x is rational' is commonly used to indicate that a certain standard has been met, and 'x is irrational' to indicate that a certain standard was not met. Furthermore, the standards are embedded in familiar contexts that are part and parcel of ordinary life (unlike, e.g., an invented game with its own standards), and rationality has to do with success in the use of some skill or ability. Correct spelling or pronunciation are similar examples. And there is no objection to claiming that they are normative concepts, if that is what the term is used to convey. But Wedgwood takes rationality to convey something more than that: That rationality is a normative concept implies "(a) rational mental states ... [and] events are in a way good, or worthy of commendation; (b) if a thinker is rationally required to ϕ then there is a sense in which the thinker rationally ought to ϕ ." (40) Now, while it follows that if one is to spell correctly one ought to conform to the standards of spelling, it is doubtful that there is a non-trivial sense of 'ought' in which it is true that one ought to follow the standards of spelling.

Does Wedgwood succeed in establishing that if one is rationally required to ϕ or to believe that p then one ought to ϕ or to believe that p? He discusses several objections, (the most important of which, the doubt about the normative implications of coherence, is addressed only in Chapter 9). According to the first objection normativity depends on how things are in the world whereas rationality is entirely mind dependent, hence they can come apart. Given that it can be bad to be rational, rational requirements are not normative. (34-5) Influenced by 'the wrong kind of reasons' writings Wedgwood suggests that the problem arises when "unusual costs and benefits, such as could be created by manipulative demons or eccentric billionaires" apply (41). In fact, such cost and benefits often arise in ordinary circumstances. For example, quite a few people may be able to obtain a job in the American Administration if they hold false views about its character. They may also do much good in the world in that case. Moreover, given what they know or are in position to know, it may be irrational of them to have those views, yet normatively speaking they should hold them, because having them is beneficial for them or for others.

Wedgwood's solution is to distinguish between two concepts of ought. He rejects the view that the distinction between right and wrong reasons is or tracks the distinction between

object and state reasons, because the latter cannot explain why whatever makes it irrational to believe a Moore-paradoxical proposition is a right reason for not believing it (45); nor can they explain why the reason not to intend to \emptyset when you have reason to \emptyset , and can \emptyset (provided that you do not intend to \emptyset), is the right reason (46). Instead he maintains that in the cases under consideration you ought to have a positive attitude toward a proposition that you ought not to have. In other words, you have a conflict precisely because you are subject to two conflicting ought propositions, though they express two different ought concepts (50). Wedgwood adds that 'the right kind of reasons' are constitutive of a kind of reasoning guiding us to an attitude or event. The wrong kind of reasons can be told apart by not being constitutive of such reasoning.

The second objection is due to the fact that it is possible rationally to believe false propositions. In such cases it is irrational to give up one's rational false beliefs, so that one is rationally required to believe a proposition that is false. If the belief is that you ought to \emptyset , you are normatively required to have it even though it is not the case that you ought to \emptyset . Hence rationality is not normative. Wedgwood rightly denies that the wide scope/narrow scope distinction refutes the objection. Instead he points out that as the rational is internal, supervening on the agent's mental states and events, the oughts that rational requirements yield are 'subjective' or information-relative oughts, relativized to the epistemic perspective of the agent who is under discussion. The subjective ought is, according to Wedgwood, relative not to all the agent's beliefs, but to a coherent epistemic perspective that he inhabits, which comprises experiences etc. besides beliefs, but is free of crazy or insane beliefs. In this he differs both from Parfit and Schroeder, who relativize to all and only beliefs, and from Kieseewetter and Lord, who relativize to available evidence only (Parfit 2011, pp. 111-113; Schroeder 2009, p. 223; Kieseewetter 2018; and Lord 2017, p. 504). While maintaining that what rationality requires is that one have beliefs required by one's relatively coherent mental states, Wedgwood allows that there may be beliefs about what would be required from another epistemic perspective; thus, he maintains, the second objection is avoided (59-60). Wedgwood allows that: It ought₁ to be that: if you believe that it ought₂ to be that you now \emptyset then you now \emptyset . If the two oughts relate to different epistemic perspectives then no contradiction need follow. Assume two propositions belonging to two perspectives: (A) you ought to believe that B; (B) you ought now to \emptyset . (The subscripts indicate that that 'ought' is perspectival.) Even if B is false, it is possible that A is true, that is, that you ought₁ to believe that you ought₂ to \emptyset now. If A and B relate to the same epistemic perspective, then Wedgwood requires that one cannot be rationally required to believe that A.

The third objection, that rationality cannot be normative because ought implies can while what we rationally ought to believe or do need not be something we can believe or do, is treated to an extended discussion (chapter 3) of the semantics of 'can' in terms of dispositions and opportunities, explained within the framework of possible world semantics. Some readers will be disappointed that Wedgwood avoided a thorough discussion of the rationale of the ought implies can slogan, and its limits, as well as the entangled questions of the dependence, if any, of cognitive states on the will.

Chapter 4 is dedicated to a criticism of a vague view called 'Reasons First', defined as holding that there is one central concept of normative reasons that is more basic and central than all other normative concepts (who belongs with this company is far from clear – I am given that honour but I deny that the concept of normative reasons is more basic than that of value). Wedgwood is familiar with a wide range of writings that he associates with this view, and he raises a number of issues they should and do contend with. Central to

those is the connection that all those writers point to (in their various ways) between the way normative reasons can guide people's reactions to the way things are, and the way they could contribute to explanations of people's reactions to the way things are. He points to a number of problems whose solution is required for 'a normative reason' to be this dual aspect concept. I found most instructive the suggestion that explanations of people's responses to reasons (as I would prefer to put the point) include factors other than beliefs, or the facts they represent. Oddly, he refrains from discussing in detail any of the solutions offered to the questions he raised, presumably thinking that as there is no common Reasons First reply none can be given – true only in that as Reasons First writers disagree, no single set of solutions would satisfy all of them.

Chapter 5 stands out in being the only one in the book offering a partial account of the meaning of an English word, 'ought'. Borrowing from possible world semantic accounts of modal concepts Wedgwood's account introduces probabilistic distributions to represent epistemic perspectives and a value function which evaluates alternatives relative to certain values. Wedgwood mentions that his preferred understanding of linguistic competence is through ability to use a word to express concepts within the range of concepts it can be used to express – a plausible suggestion which would have led him in different directions, even when dealing with the range of semantic values of 'ought'.

Chapter 6 begins with what is perhaps an over-inclusive account of virtue. Basically, if a disposition is good its goodness is a virtue, and whenever the disposition is non-accidentally manifested the manifestation has that virtue (141). So, a disposition to sneeze which is good in relieving an itch has the virtue of relieving itches, and a particular sneeze is virtuous if it non-accidentally relieves that kind of itch. "It is abstractly rational for an agent to ϕ iff there is an available way for an agent to respond to her situation that would consist of the manifestation of rational dispositions, and would result in the agent ϕ -ing. Rational dispositions are all and only those dispositions that reliably lead the agent to respond to her situation in ways in which it is abstractly rational for her to respond". (142) Disposition is a notoriously tricky concept. I suspect that Wedgwood's discussion is not helpful. Consider the following examples: I am (1) disposed to stroke my hair in a certain way, roughly meaning that when I stroke my hair, intentionally or otherwise, then other things being equal I stroke it in that way, which involves a skill to stroke it pleasantly, without scratching my skull, and without drawing attention to the activity. I am also (2) disposed to stroke it when bored. Knowing that my friends are aware of this, I am (3) disposed to intentionally stroke it when I have reason to indicate that I am bored without saying so. I am stroking my hair right now. Which, if any, of my dispositions led me to do so? Not (1): it is usually manifested when I stroke my hair but never leads to my doing so. While manifested in my stroking my hair, (3), like (1) never leads to it. (2) may lead to it, but clearly not now, as I am stroking my hair to make a point, i.e. for a reason, and none of my hair stroking dispositions led to it. Did my disposition to be rational lead to it? Assuming that I have it, it is analogous to (3) above. It is a capacity to act for a reason, which I use when I act for a reason, but it does not lead me to act for a reason. So, dispositions are involved in rational responses, in various ways, and some may lead to a rational response. But we need a finer grained analysis than is to be found here to identify them and how they do so.

As the chapter proceeds Wedgwood introduces subdivisions, which help in leading us closer to, say dispositions to behave in a virtuous way. But they do not explain the differences among dispositions, nor account for the type of dispositions that are involved in being rational. To do so we need an account of the difference between being led passively by

one's dispositions (which is all that some virtues require) and being active, as one is when acting rationally in response to reasons – these matters remain unexplored.

Chapter 7 provides an extensive defence of the internalism of rationality, as well as the claim that if it is rational for me to believe, intend etc. something then I ought (in some sense) to believe, intend, etc. it. The defence is an elaborate and careful statement of the argument from common content: "Consider two possible worlds, w_1 and w_2 . In both worlds, you have exactly the same experiences, apparent memories, and intuitions and in both worlds you go through exactly the same processes of reasoning, forming, maintaining and revising exactly the same beliefs in exactly the same ways. It seems clear that these two worlds are also exactly alike with respect to which of your beliefs are rational and which are irrational. Now suppose that in w_1 you are bedevilled by an evil demon who ensures that many of your experiences are misleading, with the result that many of the beliefs that you hold in w_1 are false. In w_2 , on the other hand, almost all your experiences are veridical, with the result that almost all the beliefs that you hold in w_2 are true. Intuitively, this makes no difference. Exactly the same beliefs are rational in both worlds." (162). Wedgwood proceeds to defend and refine this kind of argument, considering its application to narrow content in twin earth arguments and to hallucination in the case of perception. It is one of the richest and best chapters in the book, but it does leave some questions. We judge the rationality of reasoning, and other activities involving our rational powers. But we also judge the rationality of the outcomes of such activities: have we acquired irrational beliefs as a result of irrational reasoning? Etc. Wedgwood's arguments seem addressed at the activities, and leave one wondering about their results. An example: I may do something whose outcome depends on whether or not there is anything in my front garden. I believe that there is nothing, but decide, as I should, to re-examine the question. I do so, confirming my earlier belief. But my re-examination was sloppy and did not meet the standards of rationality that applied in the circumstances. My re-examination was, let us conclude, irrational. However, in the circumstances I acquired no false beliefs as a result. I believe, as I did before, that the front yard is empty. Is this belief irrational? It seems that it would be false to claim both that it is and that it is not. It is the outcome of my re-examination taken when I already believed that the yard is empty, and there is no truth of the matter whether I would have reached the same conclusion had I not that belief already (and remember that rationality is tested by recent mental events, so my belief cannot be irrational because of some event far in the past). Now change the example: during the processes of re-examination a new object was in fact placed in the yard. I would have become aware of it had the re-examination been rational. Now, it appears that my re-confirmed (and now false) belief that the yard is empty is irrational, because it is due to an irrational examination that led to a false conclusion. Wedgwood's argument refines the kind of internalism he assumes (it is required for rational requirements to be capable of guiding reasoning, i.e. transition from some mental states to others), but I think that it is at odds with, or at least undecided in, cases in which that the conclusion is false is part of what makes it irrational.

Here again Wedgwood's argument turns on the existence of dispositions to react to some mental states or events by adopting or dropping others. The dispositions are constitutive of the type of mental states involved, and the requirements of rationality express their content. As before the account strikes me as incomplete, in failing to distinguish between passive conditional reflex type dispositions and those in which the agent is active in light of his or her understanding of the reasons he or she confronts.

In suggesting that the account is incomplete I have not challenged Wedgwood's assumption

that if rational requirements are constitutive we have a disposition to comply with them. But is this a non-sequitur? An alternative would go: when thinking about how things are, for any of the many reasons that may lead one to do so or for no reason at all, one thinks in ways which subject one's thinking to the norms of that type of thought, meaning one accepts that violations of those norms would be a mistake. How does one know how to submit to norms of rationality? One learns gradually, from early childhood, as part of learning a language. Does one need to decide to use the capacity to subject one's thinking to these norms? No, their use is a skill which is automatically engaged (not always successfully) whenever one thinks about how things are. Sometimes one does not subject oneself to these norms. These are not cases in which dispositions fail. They are cases of day dreaming, fantasising, etc., rather than thinking about how things are. So understood the norms of rationality are constitutive, but do not depend on any dispositions, at least not in the way imagined by Wedgwood. But if that is the constitutive nature of norms of rationality, their normativity has to be re-examined.

For Wedgwood the route to normativity goes through the virtue of rationality, which is a special virtue, one of a "plethora of different values that are all non-relative" (i.e. do not depend on standards whose value is relative to some contingent aim) and yet the rational requirements it imposes (the ought propositions it makes true) do not conflict (as do other virtues and their requirements). Wedgwood thinks that it is in the nature of rationality that it has that character. There is no explanation of that. But there is a need to explain how such a virtue can exist. He briefly considers and rejects as inadequate to the task explanations based on Dutch book arguments and on the constitutive character of rational requirements. Rather, Wedgwood maps mental states and events present in the thinker's mind at a given time onto a rational estimate of degree of incorrectness of every way of thinking, based on the assumption that the less irrational a way of thinking is the less it is likely to be incorrect. The bulk of chapter 9 explains this system of rational estimates of the probability of incorrectness. In other words, if we can attribute a rational probabilistic assessment of the degree of incorrectness of different ways of thinking we would be able to determine the degree of rationality of different ways of thinking: the smaller the probability that they are incorrect the less irrational they are. The aim of thinking is correctness and rationality is a means to correctness, a probability-based guide to what is less likely to be incorrect. There is much to discuss regarding the presuppositions of Wedgwood's probabilistic assessment of degrees of incorrectness. But the suggestion is based on a few simpler assumptions that attract little attention in the book. In what way is rationality relevant to correctness? Some may think that it depends on the different categories of requirements of correctness. E.g. it may be a rational requirement not to believe a conjunction and the falsity of one of its conjuncts. And there may be an argument to take the correctness of a conjunction to be a reason for the correctness of each conjunct. Here a connection between rationality and correctness, though its precise form is not easy to state, seems plausible. But if my auditory experience can be a reason to believe that there is a cat around – is it a rational requirement for that belief? This is less clear in principle, and it is much more difficult to state the connection in detail. How about, my experience being a reason to believe that the person in front of me is getting irritated by my presence? It definitely may be good evidence for that belief, but what exactly is it that makes me irrational or less rational if I fail to perceive his irritation? Wedgwood's identification of rational belief with justified belief requires that not only correctness but rationality is involved in all such cases. He may be right, but the case for that view is not made in the book.

It is a virtue of the book, which is sophisticated and dense with argument about all the matters it discusses, that it ends with a list of some of the questions that remain to be explored for its account of rationality to be acceptable.